



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:17 PM UTC

PDB ID : 9HI5 / pdb\_00009hi5  
Title : Structure of scaffold FI6-focused design\_02  
Authors : Cramer, J.T.; Krey, T.  
Deposited on : 2024-11-24  
Resolution : 1.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

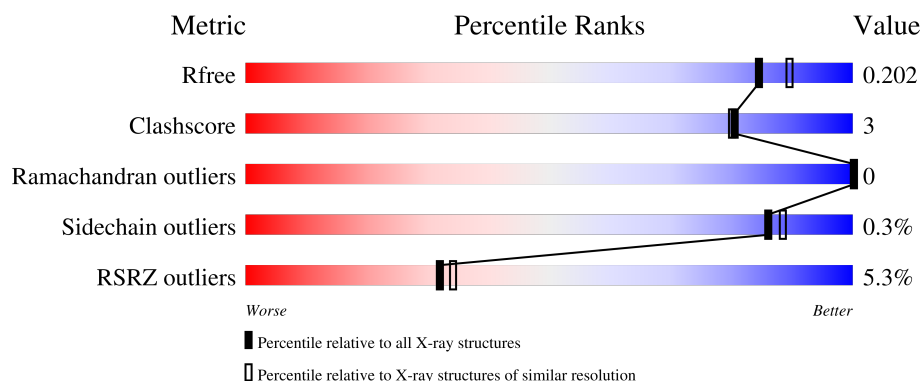
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	223	 4% 80% 6% 13%
1	B	223	 4% 95% . .
1	C	223	 4% 89% . 8%
1	D	223	 7% 83% 9% 7%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6861 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called GDSL-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	193	Total	C	N	O	S	0	2	0
			1479	945	251	279	4			
1	B	213	Total	C	N	O	S	0	0	0
			1615	1030	278	303	4			
1	C	206	Total	C	N	O	S	0	1	0
			1561	996	262	299	4			
1	D	207	Total	C	N	O	S	0	0	0
			1537	981	259	293	4			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	142	ALA	ASP	conflict	UNP A7V743
A	144	THR	PRO	conflict	UNP A7V743
A	146	ALA	LYS	conflict	UNP A7V743
A	148	ILE	MET	conflict	UNP A7V743
A	149	ASN	GLN	conflict	UNP A7V743
A	152	THR	ALA	conflict	UNP A7V743
A	155	ILE	ARG	conflict	UNP A7V743
A	156	ASN	LYS	conflict	UNP A7V743
A	159	ILE	GLN	conflict	UNP A7V743
A	165	PHE	TYR	conflict	UNP A7V743
A	169	PHE	TYR	conflict	UNP A7V743
A	170	VAL	SER	conflict	UNP A7V743
A	173	ALA	VAL	conflict	UNP A7V743
A	174	GLN	GLU	conflict	UNP A7V743
A	175	SER	GLY	conflict	UNP A7V743
A	176	PRO	ASP	conflict	UNP A7V743
A	178	GLY	LYS	conflict	UNP A7V743
A	179	ASP	ALA	conflict	UNP A7V743
A	212	GLY	-	expression tag	UNP A7V743
A	213	SER	-	expression tag	UNP A7V743
A	214	GLY	-	expression tag	UNP A7V743

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	215	SER	-	expression tag	UNP A7V743
A	216	GLY	-	expression tag	UNP A7V743
A	217	TRP	-	expression tag	UNP A7V743
A	218	HIS	-	expression tag	UNP A7V743
A	219	HIS	-	expression tag	UNP A7V743
A	220	HIS	-	expression tag	UNP A7V743
A	221	HIS	-	expression tag	UNP A7V743
A	222	HIS	-	expression tag	UNP A7V743
A	223	HIS	-	expression tag	UNP A7V743
B	142	ALA	ASP	conflict	UNP A7V743
B	144	THR	PRO	conflict	UNP A7V743
B	146	ALA	LYS	conflict	UNP A7V743
B	148	ILE	MET	conflict	UNP A7V743
B	149	ASN	GLN	conflict	UNP A7V743
B	152	THR	ALA	conflict	UNP A7V743
B	155	ILE	ARG	conflict	UNP A7V743
B	156	ASN	LYS	conflict	UNP A7V743
B	159	ILE	GLN	conflict	UNP A7V743
B	165	PHE	TYR	conflict	UNP A7V743
B	169	PHE	TYR	conflict	UNP A7V743
B	170	VAL	SER	conflict	UNP A7V743
B	173	ALA	VAL	conflict	UNP A7V743
B	174	GLN	GLU	conflict	UNP A7V743
B	175	SER	GLY	conflict	UNP A7V743
B	176	PRO	ASP	conflict	UNP A7V743
B	178	GLY	LYS	conflict	UNP A7V743
B	179	ASP	ALA	conflict	UNP A7V743
B	212	GLY	-	expression tag	UNP A7V743
B	213	SER	-	expression tag	UNP A7V743
B	214	GLY	-	expression tag	UNP A7V743
B	215	SER	-	expression tag	UNP A7V743
B	216	GLY	-	expression tag	UNP A7V743
B	217	TRP	-	expression tag	UNP A7V743
B	218	HIS	-	expression tag	UNP A7V743
B	219	HIS	-	expression tag	UNP A7V743
B	220	HIS	-	expression tag	UNP A7V743
B	221	HIS	-	expression tag	UNP A7V743
B	222	HIS	-	expression tag	UNP A7V743
B	223	HIS	-	expression tag	UNP A7V743
C	142	ALA	ASP	conflict	UNP A7V743
C	144	THR	PRO	conflict	UNP A7V743
C	146	ALA	LYS	conflict	UNP A7V743

*Continued on next page...*

*Continued from previous page...*

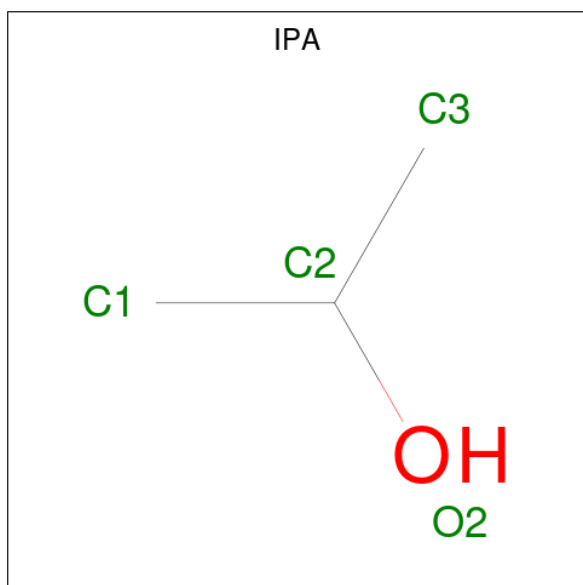
Chain	Residue	Modelled	Actual	Comment	Reference
C	148	ILE	MET	conflict	UNP A7V743
C	149	ASN	GLN	conflict	UNP A7V743
C	152	THR	ALA	conflict	UNP A7V743
C	155	ILE	ARG	conflict	UNP A7V743
C	156	ASN	LYS	conflict	UNP A7V743
C	159	ILE	GLN	conflict	UNP A7V743
C	165	PHE	TYR	conflict	UNP A7V743
C	169	PHE	TYR	conflict	UNP A7V743
C	170	VAL	SER	conflict	UNP A7V743
C	173	ALA	VAL	conflict	UNP A7V743
C	174	GLN	GLU	conflict	UNP A7V743
C	175	SER	GLY	conflict	UNP A7V743
C	176	PRO	ASP	conflict	UNP A7V743
C	178	GLY	LYS	conflict	UNP A7V743
C	179	ASP	ALA	conflict	UNP A7V743
C	212	GLY	-	expression tag	UNP A7V743
C	213	SER	-	expression tag	UNP A7V743
C	214	GLY	-	expression tag	UNP A7V743
C	215	SER	-	expression tag	UNP A7V743
C	216	GLY	-	expression tag	UNP A7V743
C	217	TRP	-	expression tag	UNP A7V743
C	218	HIS	-	expression tag	UNP A7V743
C	219	HIS	-	expression tag	UNP A7V743
C	220	HIS	-	expression tag	UNP A7V743
C	221	HIS	-	expression tag	UNP A7V743
C	222	HIS	-	expression tag	UNP A7V743
C	223	HIS	-	expression tag	UNP A7V743
D	142	ALA	ASP	conflict	UNP A7V743
D	144	THR	PRO	conflict	UNP A7V743
D	146	ALA	LYS	conflict	UNP A7V743
D	148	ILE	MET	conflict	UNP A7V743
D	149	ASN	GLN	conflict	UNP A7V743
D	152	THR	ALA	conflict	UNP A7V743
D	155	ILE	ARG	conflict	UNP A7V743
D	156	ASN	LYS	conflict	UNP A7V743
D	159	ILE	GLN	conflict	UNP A7V743
D	165	PHE	TYR	conflict	UNP A7V743
D	169	PHE	TYR	conflict	UNP A7V743
D	170	VAL	SER	conflict	UNP A7V743
D	173	ALA	VAL	conflict	UNP A7V743
D	174	GLN	GLU	conflict	UNP A7V743
D	175	SER	GLY	conflict	UNP A7V743

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	176	PRO	ASP	conflict	UNP A7V743
D	178	GLY	LYS	conflict	UNP A7V743
D	179	ASP	ALA	conflict	UNP A7V743
D	212	GLY	-	expression tag	UNP A7V743
D	213	SER	-	expression tag	UNP A7V743
D	214	GLY	-	expression tag	UNP A7V743
D	215	SER	-	expression tag	UNP A7V743
D	216	GLY	-	expression tag	UNP A7V743
D	217	TRP	-	expression tag	UNP A7V743
D	218	HIS	-	expression tag	UNP A7V743
D	219	HIS	-	expression tag	UNP A7V743
D	220	HIS	-	expression tag	UNP A7V743
D	221	HIS	-	expression tag	UNP A7V743
D	222	HIS	-	expression tag	UNP A7V743
D	223	HIS	-	expression tag	UNP A7V743

- Molecule 2 is ISOPROPYL ALCOHOL (CCD ID: IPA) (formula:  $C_3H_8O$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	3	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	161	Total	O	0	0
			161	161		

Continued on next page...

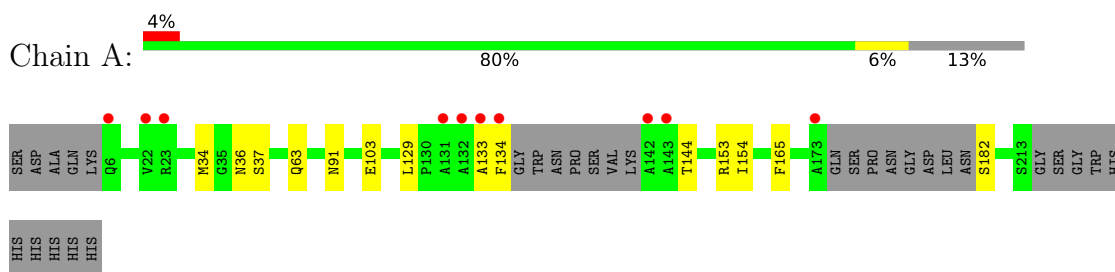
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	219	Total 219	O 219	0	0
3	C	172	Total 172	O 172	0	0
3	D	113	Total 113	O 113	0	0

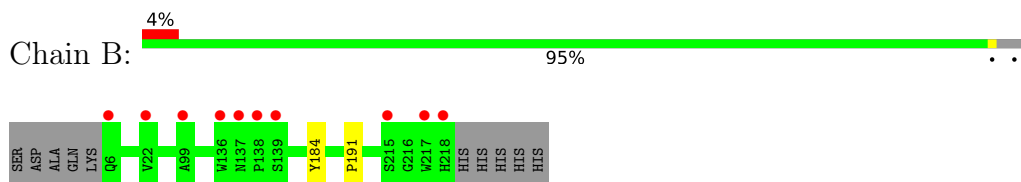
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

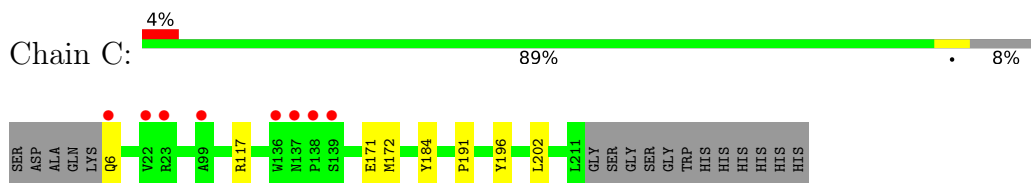
- Molecule 1: GDSL-like protein



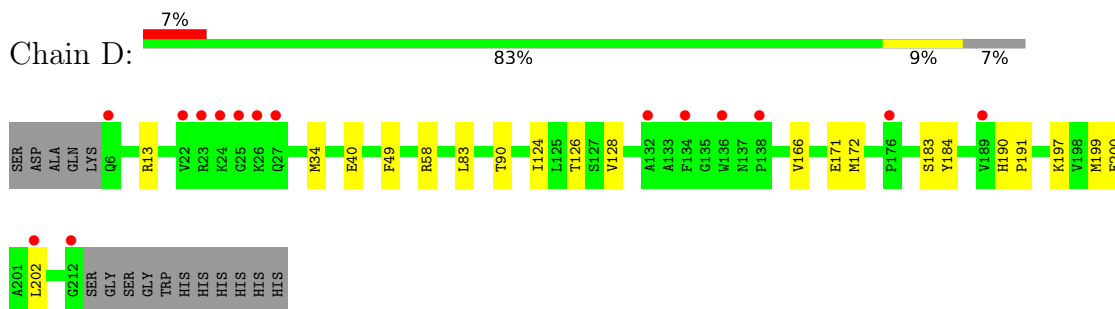
- Molecule 1: GDSL-like protein



- Molecule 1: GDSL-like protein



- Molecule 1: GDSL-like protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.05Å 63.31Å 98.18Å 96.36° 99.53° 114.57°	Depositor
Resolution (Å)	43.48 – 1.90 43.48 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.1 (43.48-1.90) 97.1 (43.48-1.90)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.91 (at 1.89Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.171 , 0.202 0.172 , 0.202	Depositor DCC
$R_{free}$ test set	3979 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.2	Xtriage
Anisotropy	0.249	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 50.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6861	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.57% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: IPA

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.47	0/1506	0.63	0/2043
1	B	0.50	0/1648	0.62	0/2240
1	C	0.39	0/1591	0.55	0/2168
1	D	0.36	0/1567	0.50	0/2136
All	All	0.44	0/6312	0.57	0/8587

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1479	0	1429	10	0
1	B	1615	0	1555	1	0
1	C	1561	0	1478	5	0
1	D	1537	0	1448	14	0
2	A	4	0	8	3	0
3	A	161	0	0	3	0
3	B	219	0	0	0	0
3	C	172	0	0	1	0
3	D	113	0	0	2	0
All	All	6861	0	5918	31	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 31 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:171:GLU:HG3	1:D:202:LEU:HD13	1.78	0.66
1:A:182:SER:N	3:A:402:HOH:O	2.29	0.65
1:D:172:MET:HE3	1:D:184:TYR:HB3	1.84	0.59
1:D:49:PHE:CE2	1:D:200:GLU:HG2	2.41	0.55
1:A:129:LEU:HD13	1:A:134:PHE:HZ	1.73	0.53

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	189/223 (85%)	181 (96%)	8 (4%)	0	100	100
1	B	211/223 (95%)	207 (98%)	4 (2%)	0	100	100
1	C	204/223 (92%)	197 (97%)	7 (3%)	0	100	100
1	D	205/223 (92%)	198 (97%)	7 (3%)	0	100	100
All	All	809/892 (91%)	783 (97%)	26 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	145/180 (81%)	145 (100%)	0	100	100
1	B	159/180 (88%)	159 (100%)	0	100	100
1	C	152/180 (84%)	150 (99%)	2 (1%)	61	61
1	D	147/180 (82%)	146 (99%)	1 (1%)	76	78
All	All	603/720 (84%)	600 (100%)	3 (0%)	86	84

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	6[A]	GLN
1	C	6[B]	GLN
1	D	183	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	45	ASN
1	D	6	GLN
1	D	91	ASN
1	D	87	ASN
1	B	145	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	IPA	A	301	-	3,3,3	0.51	0	3,3,3	0.73	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	IPA	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	193/223 (86%)	0.06	10 (5%) 33 35	14, 30, 60, 84	2 (1%)
1	B	213/223 (95%)	-0.20	10 (4%) 36 39	18, 26, 55, 84	0
1	C	206/223 (92%)	0.05	8 (3%) 43 46	20, 33, 54, 80	1 (0%)
1	D	207/223 (92%)	0.54	15 (7%) 21 23	26, 46, 73, 92	0
All	All	819/892 (91%)	0.11	43 (5%) 32 34	14, 33, 65, 92	3 (0%)

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	131	ALA	5.1
1	A	142	ALA	3.9
1	B	218	HIS	3.8
1	A	134	PHE	3.4
1	A	132	ALA	3.4

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	IPA	A	301	4/4	0.89	0.16	34,37,38,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.